

A11101 295014

NAT'L INST OF STANDARDS & TECH R.I.C.



A11101295014

/Bulletin of the Bureau of Standards
QC1 .U5 V14:1918-19 C.1 NBS-PUB-C 1905





AN "AVERAGE EYE" FOR HETEROCHROMATIC PHOTOMETRY, AND A COMPARISON OF A FLICKER AND AN EQUALITY-OF-BRIGHTNESS PHOTOMETER

By E. C. Crittenden and F. K. Richtmyer

CONTENTS

	Page
I. INTRODUCTION.....	87
II. APPARATUS.....	89
III. MEASUREMENTS ON THE IVES-KINGSBURY SOLUTIONS FOR SELECTION OF OBSERVERS.....	92
1. Preliminary measurements on solutions.....	93
2. Measurements to establish a normal characteristic ratio.....	95
IV. MEASUREMENTS ON BLUE GLASSES.....	97
V. OTHER MEASUREMENTS BY SELECTED OBSERVERS.....	104
1. On a blue solution representing the color difference of carbon and gas-filled tungsten lamps.....	104
2. On a lamp giving the color of a pentane standard against 4 wpc carbon lamps.....	106
3. On a blue solution and on multivoltage standards.....	107
VI. CONCLUSION.....	109
1. Effect of individual characteristics.....	109
2. Comparison of flicker and equality-of-brightness photometers.....	111

I. INTRODUCTION

The work to be reported in this paper was undertaken in connection with the committee on research of the Illuminating Engineering Society. The reports of that committee¹ for 1914 give a general survey of photometric problems on which investigation is especially desirable. Following this general review, it was felt that the new committee appointed for 1915 might most effectively stimulate investigation by choosing a particular field and arranging for experimental work in it. The field chosen was that of heterochromatic photometry. In order to accomplish something definite within the time available for the investigation it appeared desirable to confine the work within rather narrow limits, and it was decided to give attention primarily to the question of the methods to be

¹ Trans. Ill. Eng. Soc., 9, pp. 307, 333, 345, 358, 505; 1914.

used in actual photometric comparisons involving a color difference and to such phases of this question as could be studied in a single laboratory.

As a solution for the whole problem of photometry with a color difference the use of a flicker photometer under certain specified conditions has been proposed in particular by H. E. Ives.² To strengthen the position of the flicker instrument there has been developed also a complete scheme³ for the choice of normal groups of observers, including the establishment of an "average eye." This proposed systematization of heterochromatic measurements appeared so definite and practical as to deserve a thorough trial.

The present work, therefore, was planned to show the difference to be expected between individuals and to include readings by a large number of observers so as to establish average or normal values for various measurements involving color differences. In general, similar measurements were to be made on a flicker photometer and on an equality-of-brightness photometer in order to establish the relation between results obtained by the two methods and the relative certainty of measurements made by the two types of instruments.

The experimental data to be presented were obtained in the laboratories of the Bureau of Standards during the summer of 1915. Besides extensive preliminary tests, the data obtained include (1) readings by 115 observers on the Ives-Kingsbury test solutions for choice of observers, (2) measurements by the same observers on blue glasses presenting a color difference equivalent to that involved in comparing carbon lamps with vacuum tungsten lamps, (3) a repetition of the above measurements by a selected group of observers, (4) sets on a blue solution corresponding to the color difference between a carbon lamp and a gas-filled tungsten lamp, (5) a direct comparison of lamps operated at the color of the pentane lamp flame with others run at 4 watts per candle, and (6) the calibration of a blue solution and measurements with it on lamps at various efficiencies.

With the exception of the solution used in testing observers, it may be noted that the work has dealt only with color differences of the type given by two incandescent bodies at different temperatures, such as two lamps operated at different efficiencies. Lights showing this type of color difference are, of course, much more

² *Phil. Mag.* (6), 24, p. 852, 1912; *Trans. Ill. Eng. Soc.*, 10, p. 317, 1915.

³ *Ives and Kingsbury, Trans. Ill. Eng. Soc.*, 10, p. 203; 1915.

easily compared than those showing a more nearly "saturated" hue, but the difficulties are sufficient to impair very seriously the accuracy of many practical photometric measurements required at the present day. It is highly desirable that a method of comparing the intensities of lights of different colors which can be used for all types of color difference should be agreed upon, but at present the field in which there is most urgent need of a high degree of accuracy in such comparisons is the rating of incandescent lamps. In the present investigation it has appeared desirable to make those tests which would have the most direct bearing upon the practical application of the instruments and methods involved.

II. APPARATUS

Two standard photometer bars were arranged as nearly as possible alike, on one of which a flicker photometer was used and on the other a Lummer-Brodhun photometer. In each case the photometer head was stationary and was illuminated on the left by a stationary lamp placed at such a distance as to give an effective brightness of about 2.5 millilamberts in the photometric field after allowing for all losses in the apparatus. A similar lamp on a carriage at the right was moved by turning a wheel beneath the photometer, and the cells and glasses referred to later were inserted on this side of the photometer so that a constant illumination was maintained. Settings of the movable lamp were printed on a record sheet and were measured from reference lines on the sheet, proper allowance being made for the optical thickness of the cells. This method of recording settings is much quicker than reading from the bar; it also has the advantage of giving a permanent record free from the errors which are likely to be made in transcribing numerical readings. The mean of the groups of points can be located very quickly with a sufficient degree of accuracy. In these tests each observer was asked to keep a tally of the settings, which required taking the hand from the wheel which moved the lamp, as well as turning away from the photometer.

The lamps used have double hairpin carbon filaments in one plane, and were operated at a voltage which made them match the color of the Bureau's 4 wpc standards. For the distances at which the lamps were used, the illumination given follows the inverse-square law with a sufficient degree of exactness so that no corrections were necessary. The voltage was controlled by potentiometers.

The flicker photometer used was a standard Lummer-Brodhun head with the rotating prism attachment described by E. F. Kingsbury.⁴ The particular instrument used was very kindly loaned by Mr. Kingsbury. The photometer head was modified by removing the original prisms and putting in a pair of which one has two quadrants cut away so as to make the duration of exposure to each light the same. (See Fig. 1, A.) In the flicker attachment as originally made the focal plane of the eyepiece fell considerably beyond the comparison prisms. By inserting a collar to extend the telescope the instrument could be made to focus on the face of the prisms, this arrangement being intended for use in making equality-of-brightness settings with the flicker prism at rest. A considerable number of trials indicated that for

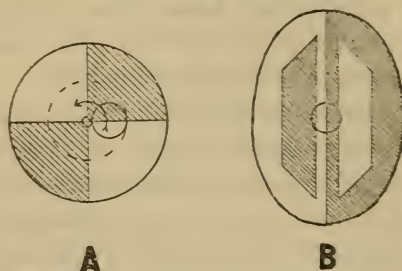


FIG. 1.—Photometric fields

A. Type of field used in the flicker photometer. The small circle which one observes travels over the field as indicated. B. Lummer-Brodhun field used in the equality-of-brightness measurements. (Contrast type, not indicated by sketch.) The circle shows the part used for the small field, the remainder being covered by an illuminated diaphragm

most observers this arrangement was decidedly better for flicker settings. Consequently it was used throughout the tests. Without a very good photometric field this could not be done, since any imperfections in the field would cause flicker, but the prism used was sufficiently good so that with a color and intensity match there was practically no flicker in the field even at low speeds.

A tachometer was attached to the flicker mechanism, and the speed was controlled by a rheostat in series with the motor. It was expected that each observer would have to choose a suitable speed for each color difference measured, but extensive trials with a number of observers showed that, over a limited range, change of speed had very little effect either on precision of setting or on the mean result. It was finally decided to adopt a moderate speed (12 light cycles per second) for all observers and all set-

⁴ Jour. Franklin Institute, 180, p. 215; 1915.

tings. Observers were then directed to set for a minimum of flicker. Although this speed was rather low for settings on the yellow solution and high for those with color match, few observers found serious difficulty in making definite settings. The constant speed was adhered to partly because in the preliminary measurements there had been indications of slight changes of results when speeds much higher or much lower were used, particularly with the yellow solution.

On the second photometer bar measurements were made under two conditions—(1) with the standard Lummer-Brodhun contrast field, (2) with an illuminated diaphragm which limited the field to about 2° , as in the flicker instrument. (See Fig. 1, B.) In the case of the contrast field, however, all observers were asked to make their settings by the middle strips, disregarding the contrast trapezoids. The small field was used because it gave equality-of-brightness measurements under conditions closely similar to those used with the flicker instrument. It was also thought that limiting the field so as to use only a fairly homogeneous part of the retina might reduce variations in judgment and give settings more truly characteristic of the observer's eye.

All the photometers used had pupil apertures 5 mm in diameter. Each photometer head was provided with holders for absorption glasses, as well as holders for absorption cells, especially constructed to prevent any diffusion of stray light into the photometric field.

Three pairs of 1 cm cells were provided. These were constructed practically like those described by Ives and Kingsbury,⁵ with removable sides of colorless optical glass. A small modification which was found to facilitate secure sealing of the sides was made by beveling slightly each edge of the cell blocks, thus making a groove to be filled by the paraffin seal. Although the sides were made of polished plate glass, it was found that there were appreciable differences between the transmission of different plates. Repolishing of the plates, which was necessary after using some solutions, also changed the transmissions perceptibly. The differences and the changes mentioned were less than 1 per cent, but were not negligible for precise work. So far as color is concerned, the glass used was satisfactory. It had been obtained for other work, requiring very clear glass, and its applicability for the present purpose was tested directly by measuring its transmission for

⁵ Trans. Ill. Eng. Soc., 9, p. 795; 1914.

the two extreme colors of light to be used—that is, the light transmitted by the two test solutions. Its transmission for the two was the same within 0.2 per cent.

Considerable time was given to experiments with the cells to determine how closely conditions could be reproduced. In brief, it may be said that results can be rather easily reproduced to within 1 per cent with them, but if an accuracy greater than one-half per cent is desired, extreme care is necessary. In any case thorough cleaning is essential, and it is desirable to compare the cells with each other after each cleaning.

The paraffin used for sealing the cells is somewhat difficult to remove completely from the plates, and a supply of hot, running water is almost a necessity for thorough cleansing. Xylol also is a convenient solvent of paraffin and is especially useful when the supply of hot water is not plentiful.

Cells should be filled well up to the neck, for if a meniscus of considerable size is left an appreciable amount of light may be reflected from it into the field.

III. MEASUREMENTS ON THE IVES-KINGSBURY SOLUTIONS FOR SELECTION OF OBSERVERS

The method of selecting observers, to which reference has been made,⁶ is based on determinations of the relative transmission of two solutions—one reddish yellow, the other blue-green. These are solutions, in water, of potassium bichromate and of copper sulphate, containing, respectively, 72 and 53 grams of the salt per liter of solution. While not definitely specified in the original proposal, it has been assumed that the solution is to be made up at 20° C, and that by copper sulphate is meant the crystals $\text{CuSO}_4 + 5\text{H}_2\text{O}$.

When measured at 20° C by the “average eye” with a flicker photometer conforming to specifications previously mentioned, 1 cm layers of these two solutions were intended to have equal transmissions for the light of a carbon lamp of the standard 4-watt-per-candle color. The average eye thus defined was originally established by measurements made by 61 observers on the transmission of a green solution.⁷ The two solutions above described were worked out later on the basis of measurements made by selected groups of observers. It is not at all clear that an average established by comparing the middle of the spectrum

⁶ *Trans. Ill. Eng. Soc.*, 10, pp. 203-208; 1915. ⁷ *Ives and Kingsbury, Phys. Rev.* (2), 5, p. 230; 1915.

with the whole spectrum can be legitimately thus transferred by a few observers to the basis of a comparison of the two ends of the spectrum. For instance, it is known that an observer may be abnormally sensitive or nonsensitive in the middle of the spectrum and yet appear normal in the comparison of the two halves of the spectrum, or he may be normal according to the first test and not so by the second. In fact, it would appear that tests of both kinds should be included in choosing observers for measurements of illuminants which show marked selectivity in the visible spectrum. Ives and Kingsbury state, however, that groups of observers selected by one criterion were found to satisfy the other. For the types of color difference with which the present investigation has been most directly concerned the two-solution test appeared most significant, besides being more convenient than the earlier one. This method alone, therefore, has been used for testing observers, and the characteristics of a given observer will be supposed to be represented by the ratio of the transmission of the yellow solution to that of the blue solution, as measured by that observer, although it is recognized that this ratio is more strictly an index of the observer's sensitiveness to lights in which different proportions are contributed by the two ends of the spectrum.

1. PRELIMINARY MEASUREMENTS ON SOLUTIONS

The standard temperature for the solutions is 20° C. The greater part of the present work was done at temperatures ranging from 25° to 30°, and, consequently, it was necessary to determine the temperature coefficients of the transmission of the test solutions. Over the range considered the variation with temperature was found to be practically linear. The transmission of the potassium-bichromate solution decreased nearly 0.2 per cent per degree rise of temperature, while that of the copper-sulphate solution decreased about half as much. The differential correction to be applied to the ratio of the two transmissions was, therefore, practically 0.1 per cent per degree centigrade, the observed ratio (Y/B) being too small when the temperature was above 20°.

Other conditions which affect the value of the ratio obtained are the color of the light for which the transmissions are measured and the brightness of the photometric field. The color is supposed to be that of the standard 4 wpc carbon lamp and the effective brightness 2.5 millilamberts (equivalent to an illumination of 25

meter-candles on a perfectly diffusing and completely reflecting surface) after allowing for losses in the photometer. In the particular instruments used these losses aggregated nearly 50 per cent, so that the actual illumination necessary was about 50 meter-candles.

On account of the lack of a convenient nomenclature, as well as the difficulty of absolute determinations of diffuse reflectivities of surfaces, there has been some confusion as to the exact field brightness used in various investigations. In making the measurements recorded in this paper the effective brightness used was very close to 2.5 millilamberts; that is, the brightness produced by an illumination of 25 meter-candles on a perfectly diffusing and completely reflecting surface. The numerical values for the test ratios originally given before the Illuminating Engineering Society⁸ were reduced, however, to the basis of an illumination of 25 meter-candles on a white surface having a reflectivity of approximately 90 per cent. Since a specification of the absolute brightness is preferable, 2.5 millilamberts is here retained as the standard brightness, and test ratios are given in this paper on that basis. The difference is quite unimportant, since the correction previously applied to the ratios was only 0.003.⁹

Since the variation of results arising from changes in the efficiency of the lamp or in the illumination used is small, no very precise determination of the effects of such changes has been made. Some measurements were made, however, with a lamp operated at 3.1 and at 5 watts per candle and with effective illuminations of approximately 10 and 50 meter-candles. In accordance with the reversed Purkinje phenomenon shown by the flicker photometer,¹⁰ it was found that the ratio of transmissions (yellow ÷ blue) was smaller at the higher illuminations. At 50 mc the average of three observers gave a ratio slightly over 1 per cent lower than the normal, while at 10 mc the ratio was 2 per cent higher than normal. A rise in the efficiency of the lamp naturally causes a decrease in the observed ratio, but the variation is so small that the effect of any error likely to occur in the rating of the lamp would be entirely negligible. Running the lamp at 3.1 wpc, or at 5 wpc, instead of 4, causes a departure from the normal ratio of only 1 to 2 per cent. It may be well to record that the funda-

⁸ Trans. Ill. Eng. Soc., 11, p. 331 (also note, p. 333); 1916.

⁹ Thus, in the earlier paper the average observed ratio 0.987 was "corrected" to 0.990, whereas in either case the round number 0.99 may as well be used for all practical purposes. Fig. 6, in which the change would be scarcely perceptible, has not been redrawn.

¹⁰ Ives, Trans. Ill. Eng. Soc., 5, p. 717, 1910; Phil. Mag. (6), 24, p. 170, 1912.

mental 4 wpc carbon standards have oval-anchored filaments, and since their average reduction factor is 0.825 the standard efficiency is 4.85 watts per spherical candle or 2.6 lumens per watt. The color of the light is practically the same as that given by a vacuum tungsten lamp at 3.1 wpc or 3.2 lumens per watt.

The composition of the test solutions is such that there is no reason to expect any difficulty in reproducing them or any change with time. Cells used for several months during the investigation showed no appreciable change in transmission.

2. MEASUREMENTS TO ESTABLISH A NORMAL CHARACTERISTIC RATIO

In order to obtain an independent check on the average eye as defined by the test solutions and to test the characteristics of observers to be used in later measurements, and at the same time to establish the relation between the characteristic ratio and some measurement met with in practical work, a series of measurements was made by 115 observers. On each photometer this series consisted of 8 sets of 10 readings each. On the flicker photometer sets were made on the two lamps at color match and on the same lamps with a blue-glass screen and with cells containing the two test solutions interposed in succession on one side of the photometer head. The four sets were immediately repeated in different order. The data given, therefore, are the mean results of 20 settings of the photometer on each condition.

The original plans were to include a similar series of measurements on the equality-of-brightness photometer, but with it only the most experienced observers could make any definite settings on the test solutions, and even they varied so greatly from day to day that the results were quite useless as an indication of the observer's color characteristics. For instance, ratios determined on one day by the mean of several hundred settings could not be repeated within 10 per cent on the next day, although the same observers could reproduce their ratios day after day on the flicker photometer with an average deviation of less than 1 per cent, only 20 settings being made on each solution. Consequently, the equality measurements were made only on the smaller color differences, such as that presented by the blue-glass screen, for which the results are given in the next section. One observer was unable even then to make settings definite enough to be used, and his results are consequently not included in the following data, although his flicker settings were good.

With very few exceptions the observers included in these measurements are men who have had some years of experience in physical or chemical observations, and nearly 30 of them have had considerable recent practice in some sort of photometric measurements. The distribution of 114 observers with respect to characteristic ratio ($Y \div B$) is shown in Fig. 2, where the ordinates represent the number of individuals falling within a range of 1 per cent. For example, between 0.900 and 0.909, inclusive, there are 5 observers. Such frequency curves must be used with caution,

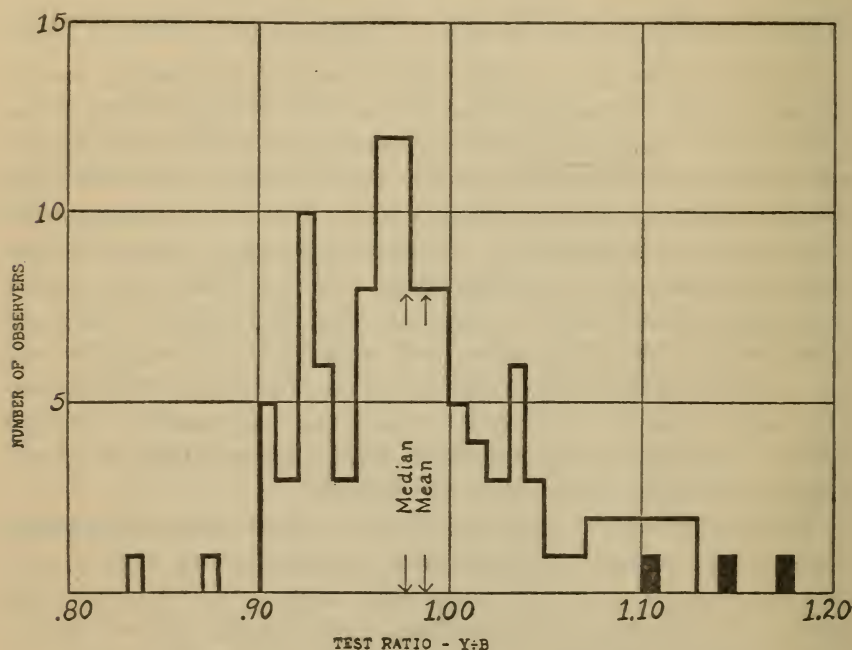


FIG. 2.—Distribution of 114 observers with respect to characteristic ratio (ratio of transmission of yellow test solution to that of the blue solution)

The ordinates show the number of observers falling in a range of 0.01 in ratio. The black rectangles represent "color blind" observers

since their shape can often be greatly changed by grouping in different ways. In this case the number of individuals must be considerably increased before it can be told with certainty whether the unsymmetrical shape of the curve is accidental, but there is a marked indication of the existence of a fairly definite type of eye which gives systematically high ratios in the neighborhood of 1.10 to 1.12. The solid black rectangles represent three men who were found to be definitely color blind in a test made by I. G. Priest with Nagel test cards. Some others in this group are known to have

peculiarities in their color perception, although they are able to pass color-vision tests.

The arithmetical average of the 114 ratios is 0.987. If the three observers classed as color blind are omitted, the average is 0.983, but actually the line between normal and abnormal color vision is very difficult to draw. Moreover, the relation between color and luminosity is not very definite, and in making up an average luminosity scale there is little justification for ignoring that percentage of the people who are color blind. On the other hand, taking an average may give undue weight to the abnormal observers, and there is some advantage in taking instead of the average the median value; that is, a value such that there are equal numbers of observers above and below it. In this case the median for the 114 ratios is 0.977, and the effect of omitting the three extreme observers mentioned is only to make the median fall between 0.976 and 0.977. To show how closely groups selected at random might be expected to agree on the ratio, these 114 observers were arranged alphabetically and for each half of the list the mean and the median values were found. The two means were 0.982 and 0.992; that is, 1 per cent different. The two medians were 0.975 and 0.979, or 0.4 per cent different. In other words, in reproducibility the median appears to be somewhat better than the mean. Even if the three color-blind observers are omitted, the means of the two groups still differ by 0.8 per cent, being 0.979 and 0.987.

In order to test the constancy of the characteristic ratio of individuals, 20 observers repeated their measurements at the end of the test. The average deviation of an individual from his first value was 1 per cent, and the mean of the 20 differed by 0.2 per cent in the two sets. Repeated measurements during the several months occupied by other parts of the work have indicated that this is the amount of variation to be expected in successive sets. Consequently, the average deviation of an experienced observer from his mean value is usually well below 1 per cent. No definite indications of any important change in an individual's ratio have been found.

IV. MEASUREMENTS ON BLUE GLASSES

The blue glasses mentioned presented a color difference equivalent to that between a 4 wpc carbon lamp and a vacuum tungsten lamp at about 1.2 wpc (8.2 lumens per watt). In other words,

when the glass was placed in front of a carbon lamp the light transmitted was similar in color to that from a tungsten lamp, and it was compared with unmodified light from another carbon lamp. On such glasses measurements were made by all the observers with the flicker photometer and with the two forms of equality field. The latter both showed large variations in results, and the general result is perhaps better shown by averaging out some of the individual errors. In order to do this the observers

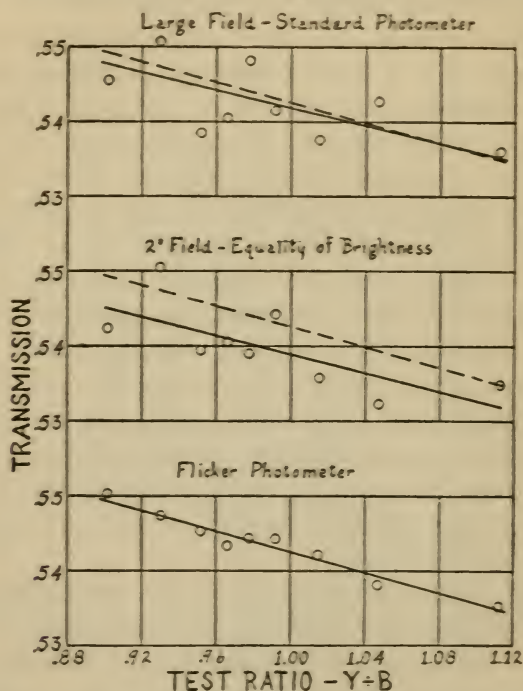


FIG. 3.—Relation between observed transmission of blue glass (3G) and characteristic ratio of observer

The color difference involved is equivalent to that between 4 wpc carbon lamps and 1.1 wpc tungsten. Each point is the mean of 12 or 13 observers. The solid lines represent the least square solutions for 114 observers. The flicker photometer curve is drawn (dashed) with the others for comparison.

were arranged in the order of their characteristic ratios and averaged in groups. Three of the extreme groups consist of 12 observers each, the others of 13. The results are shown in Fig. 3. The flicker photometer data are plotted at the bottom, and for comparison the curve drawn to represent them is reproduced in a broken line along with the data for the two equality fields. The curves drawn represent the least square solutions for the whole 114 observers, assuming a linear relation between characteristic

ratios and observed transmissions. In the flicker measurements the change in transmission corresponding to 1 per cent difference in ratio is 0.13 per cent, while the other curves show corresponding changes of 0.11 and 0.12 per cent.

The observers were somewhat arbitrarily classified by inspection of their settings on the equality photometer with regard to the consistency of settings (not their accuracy), and in Table 1 are given the mean results of the three classes, "a" meaning good, "b" medium, and "c" poor sets. The transmissions are all reduced to the basis of the mean ratio (0.987) so as to be strictly comparable, and the residuals are departures from the curves of Fig. 3; that is, the systematic errors due to individual characteristics have so far as possible been eliminated and the residuals represent largely the accidental errors in judgment. All transmissions given in this paper are for light of the quality given by a 4 wpc carbon lamp.

TABLE 1
Transmission of Blue Glass (3G) and Residual Errors (114 Observers)

Class	Number of observers	Transmission			Mean residuals (per cent)		
		Flicker photometer	Equality photometer		Flicker photometer	Equality photometer	
			Large field	Small field		Large field	Small field
a.....	31	0.5430	0.5426	0.5404	0.5	1.2	1.2
b.....	58	.5436	.5429	.5369	.6	1.7	1.5
c.....	25	.5437	.5414	.5454	.7	3.2	3.5
Mean.....	114	.5434	.5425	.5396	.6	1.9	1.9

To test the reproducibility of results 20 observers well distributed with respect to ratio, 12 of whom had had considerable photometric experience, were selected to repeat these measurements. The individual observations are shown in Fig. 4, in which, as before, the flicker curve is drawn through the other data for comparison. The average values for the transmission are given in Table 2.

TABLE 2

Transmission of Blue Glass (3G)—Two Sets by 20 Observers

	Flicker photometer	Equality photometer	
		Large field	Small field
First set.....	0.5429	0.5431	0.5447
Second set.....	.5422	.5472	.5456
Mean.....	.5426	.5452	.5452
Average difference between two sets by each observer (per cent)....	.4	1.4	1.2
Mean obtained by 14 more consistent observers.....	.5427	.5472	.5436

Both groups of better observers in Table 1 find the transmission of the blue glass smaller with the small field than with the large,

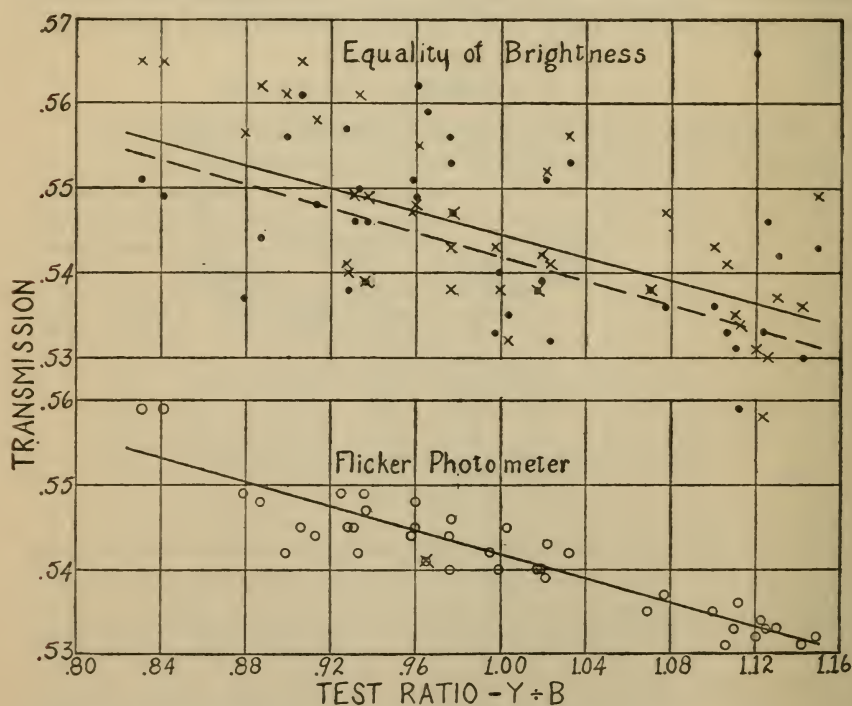


FIG. 4.—Transmission of blue glass (3G)—two sets by each of 20 observers

The crosses represent measurements with a standard Lummer-Brodhun photometer, the dots those with a small field. For this group of measurements the average results with these two fields were the same. (See Table 2.) The dashed line is the flicker curve

which is to be expected, but obtain higher values with the flicker than with either form of equality field. The 20 who were chosen to repeat the measurement happen, however, to read, on the aver-

age, higher on the equality fields. The number of sets made is by no means sufficient to establish standard values with the equality photometer, and a better approximation to the real "average" transmission can perhaps be obtained by omitting the observers whose sets were not consistent. Six of the 20 observers in the second set obtained results on one or both of the equality fields which differed by 2 per cent or more from their first sets. If these six are omitted, the means of the remaining 28 measurements are as shown in the last line of Table 2.

With practice most observers tend to fix upon a more definite concept with regard to what constitutes equality of brightness between two different colors, and the question has been asked whether experienced observers obtain the same result as inexperienced ones. Naturally the practiced observer can make more precise settings. In other words, to reproduce results with a given percentage of accuracy the unpracticed observer must make more measurements, but it does not appear that this fact gives any reason for expecting a systematic difference between results obtained by experienced observers and those obtained by inexperienced persons. However, of these 114 observers the ones who might be classed as experienced in photometry did get a result slightly higher than the others. Twenty-two observers were selected as having had such experience as would justify giving extra weight to their sets if the purpose of the paper was to establish normal values for the Lummer-Brodhun photometer, and their results follow in comparison with the mean of all the observers.

TABLE 3
Transmission of Blue Glass (3G)

Observers	Flicker photometer	Equality photometers	
		Large field	Small field
114	0.5434	0.5425	0.5396
22 "experienced"5437	.5462	.5434

Among these 22 experienced observers are included half of the group whose results are given in Table 2, but the agreement between the two groups would not be materially affected if the observers common to the two were omitted.

Table 1 by itself would indicate a high degree of certainty in the transmission as determined by the large-field equality photometer,

but this certainty is reduced by the failure of the 20 observers to repeat their results or to agree with the larger group. Nearly half of the change shown by the 20 observers is due to one very poor set, but of these 20 observers the 14 who repeated results most consistently got a value 0.7 per cent above the mean of all the 134 observations, and this higher value is corroborated by the other experienced observers.

As the result of the 134 measurements, the unweighted average values for the transmission of this glass are 0.543 for both flicker and large-field equality and 0.541 for the small-field equality photometer, but in general the observations here recorded leave an uncertainty of the order of 1 per cent in the values to be assigned for measurements by the equality photometer. Further measurements, discussed in a later section, indicate that the values obtained by the more consistent and the more experienced observers (Tables 2 and 3) are nearer the result which would be obtained by increasing greatly the number of observations. The 22 observers whose Lummer-Brodhun settings should have most weight on the score of experience obtain practically the same result with the two small fields (flicker and equality) and about one-half per cent higher with the large field. In searching for such small differences, however, it is sometimes misleading to take averages of a small number of observers without considering the individual observations, for if most of the observers get small differences, a few erratic observers may throw the average results to one side or the other. Comparing the two small fields of these 22 observers 11 are higher on the flicker and 11 higher on the equality. As between the large-field equality and the flicker, 13 read higher on the former and 9 higher on the latter. While there is a definite preponderance in the one direction, it is evident that the magnitude of the difference is not established with much accuracy.

Middlekauff and Skogland¹¹ have already called attention to the relation between the above values and those obtained in several other laboratories. This glass has a transmission 1.6 per cent greater than the glass 3B included in their comparative measurements, and if the values assigned for 3B are correspondingly increased for comparison with the preceding (0.543), the results in different laboratories are 0.538 with a flicker photometer and 0.546, 0.551, and 0.552 with Lummer-Brodhun photometers.

¹¹ *Trans. Illum. Eng. Soc.*, 11, p. 164, 1916; this *Bulletin*, 13, p. 287, *Sci. Paper No.* 277.

There are two differences in conditions which may in part account for the fact that the result obtained in the present work is below all the other Lummer-Brödhun values. In this work the photometer was used as an "equality" rather than a "contrast" field, and the illumination was much higher than that used in the other measurements. The difference probably arises, however, more from the fundamental uncertainty of equality-of-brightness measurements than from any of these systematic differences in conditions.

Whether or not there is a systematic difference between the results obtained by contrast settings and those given by strictly equality settings can not be told except by testing a large number of observers, for on changing the method of judgment each observer must to a considerable extent reestablish his "definite concept" of what constitutes a setting, and some observers change in one direction, some in the other. In general, the mean result obtained by a group of observers does not seem to be changed definitely in either direction by changing the method of judgment. When there is a considerable color difference in the field, some observers can set more definitely if the contrast strips are removed. In this case the strips were not removed, but observers were asked to disregard them so far as possible and to set for equality of brightness in the central strips of the field. It is to be noted that the 22 "experienced" observers setting this way obtain exactly the same average result as Middlekauff and Skogland's group setting by contrast.

The greater certainty of the flicker values is shown by a comparison of the mean residuals in Table 1 and the differences between sets in Table 2. It may be remarked that some observers with practice develop the ability to repeat values very closely on the equality photometer, and such observers would make a much better showing for that photometer in a comparison like that of Table 2. Unfortunately, however, such observers are comparatively rare and do not in all cases settle on a value in agreement with their characteristics as indicated by the flicker method.

Of the equality-of-brightness photometers the small field shows no material superiority over the large one. Table 2 would seem to indicate that it gave more reproducible results, but in Table 1 it will be seen that the residuals average the same for the two forms, while the small field shows the poorest agreement between groups of observers. The use of the small field was discontinued after this test, because it gave no promise of any practical advan-

tage. The data obtained are sufficient to indicate that if the small field had been carried through the measurements on the larger color differences the results would almost certainly have agreed with the flicker values more closely than the large-field measurements did. Practically, however, with these larger differences equality-of-brightness settings show such large variations from time to time and such wide differences between individuals that the results have very little significance. The use of the large field was carried as far as practicable because the tendency has been to use that type of photometer for all sorts of measurements, and it was desired to correlate that instrument with the flicker, particularly with respect to differences between individuals as well as relative average values for the two instruments.

V. OTHER MEASUREMENTS BY SELECTED OBSERVERS

The preceding sections cover the primary purposes of this work, which was to establish an average eye and to test the usefulness of the proposed method of selecting observers by having a large number make measurements on the test solutions and on some typical color difference. The following sections give the results of measurements by a few observers with several other degrees of color difference. The observers were selected from those conveniently available, the principal purpose being to secure a wide range of individual characteristics in order to illustrate the application of the method of selection. Only a few measurements were made, and the results are intended to be qualitative demonstrations of the method, not precise determinations of the things measured.

1. MEASUREMENTS ON A BLUE SOLUTION REPRESENTING THE COLOR DIFFERENCE OF CARBON AND GAS-FILLED TUNGSTEN LAMPS

This color difference is so great that most observers showed marked fluctuation in results from day to day with the equality photometer, although the majority made fairly good settings at any one time. Eleven observers made two sets on both photometers, using only the large field on the equality of brightness. On it each set consisted of three separate groups of readings, while on the flicker only two groups were made for each set, but the equality results were so erratic that the whole series of measurements with it were repeated. The individual sets are shown in Fig. 5. The circles crossed by a line indicate sets in which means of groups showed marked discrepancies among themselves. The mean of the equality-of-brightness sets is about 5 per cent above

the flicker curve, but the only observers who repeated values at all consistently from day to day fall closer to that curve, and the most consistent observer has all four sets below it. Nevertheless, even if the more erratic sets are discarded, the equality values are definitely higher than those obtained by the flicker method.

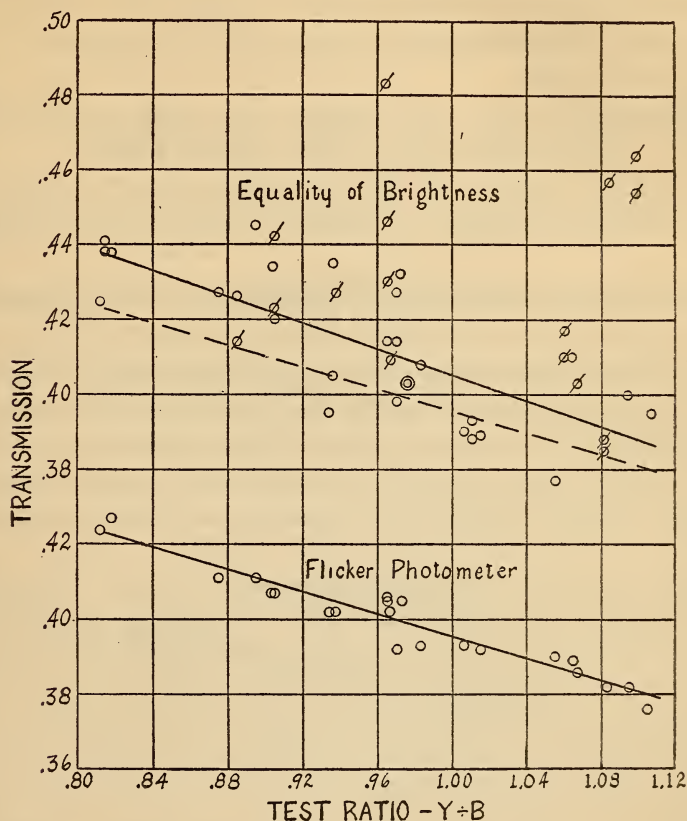


FIG. 5.—Transmission of 1 cm cell of 75 per cent concentration Ives-Kingsbury nickel-ammonium sulphate solution

The crossed circles indicate measurements in which the groups of readings differed greatly. These are disregarded in drawing the curve. The published equation for transmission of this solution gives a value of 0.397 for this concentration

For a ratio of 0.99 the mean transmission by the flicker photometer is 0.397; by the equality (including all sets) 0.417, the mean residuals being 0.8 per cent and 3.5 per cent. The slope of the flicker curve indicates a change in transmission of 0.38 per cent for 1 per cent difference in ratio.

It may be noted that this solution is the 75 per cent concentration of Ives and Kingsbury's blue working solution,¹² for which

¹² Trans. Ill. Eng. Soc., 10, p. 253; 1915.

the transmission calculated from the published equation is 0.397, as determined for a characteristic ratio of unity, which should give 0.399 for a ratio of 0.99. The difference between the original calibration and the present check on this one point with the flicker photometer is therefore one-half per cent. There are no other equality-of-brightness measurements available for comparison. Some further measurements on this solution are given in a later section of this paper.

2. MEASUREMENTS ON A LAMP GIVING THE COLOR OF A PENTANE STANDARD AGAINST 4 WPC CARBON LAMPS

Fig. 6 shows similar data for the comparison of a carbon lamp with a standard operated to match the pentane lamp flame in

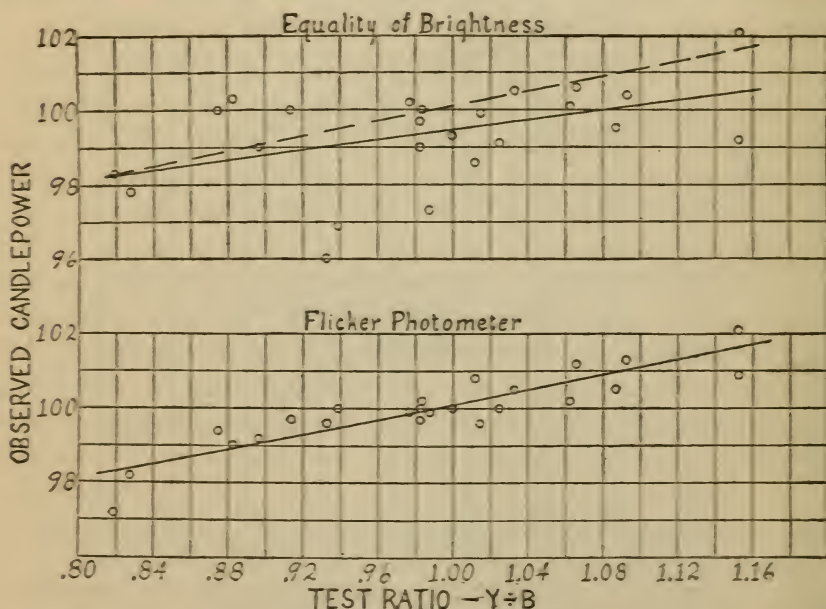


FIG. 6.—Measurements of the candlepower of a lamp matching the pentane flame in color, using a 4 wpc carbon standard

The candlepower scale is chosen to make the average flicker photometer value 100

color. This is in the neighborhood of 7.5 wpc. Here the slope is reversed because the light measured is redder than that of the carbon lamp, and the change in observed candlepower corresponding to 1 per cent in ratio is 0.1 per cent.

Each observer made two sets on each photometer, a set on the equality consisting of three groups of readings, but on the flicker of only two groups. The average difference between the two sets on the flicker was 0.6 per cent, on the equality 1.2 per cent, the

average residuals (departures from the curve) being 0.4 per cent and 0.9 per cent. In the two measurements of the characteristic ratio made by each observer the average difference was as usual 1 per cent.

The mean value obtained on the equality is 0.6 per cent below that given by the flicker photometer, and that this is not entirely due to accidental variation is indicated by the fact that 10 of the 12 observers varied in this direction.

The actual candlepowers found were 8.75 by the flicker and 8.70 by the equality photometer for a lamp whose value had previously been established as 8.70 by repeated calibrations in comparison with 4 wpc standards on the standard Lummer-Brodhun photometer.

3. MEASUREMENTS ON A BLUE SOLUTION AND ON MULTIVOLTAGE STANDARDS

At the completion of this work the lamps which had been measured at various laboratories, as reported by Middlekauff and Skogland,¹³ were available, and in order to obtain a direct comparison of further flicker values with the results of those measurements a few sets were made on the lamps at different voltages. As a convenient means of making the measurements with the standard illumination on the flicker photometer, the values for the lamps were obtained indirectly by first calibrating, with the flicker photometer, cells filled with properly chosen concentrations of the Ives-Kingsbury blue solution, which has already been mentioned, and then measuring the lamp on the Lummer-Brodhun photometer with approximate color match obtained by the cells. This procedure also gave a check on the calibration of that solution as published. The measurements made on each concentration of the solution were four sets by each of three observers, with slight corrections to put the results on the basis of the average eye giving a characteristic ratio of 0.99. The results are given in Table 4, the transmissions being percentages of the transmission of a similar cell filled with water. The values in column 4 are calculated from Ives and Kingsbury's equation. These values for transmission all refer to the transmission for light similar in quality to that of a 4 wpc carbon lamp.

¹³ Trans. Ill. Eng. Soc., 11, p. 164, 1916; this Bulletin, 13, p. 287, Sci. Paper No. 277.

TABLE 4
Transmission of 1 cm Cells with Ives-Kingsbury Blue Solution

Equivalent watts per candle	Concentration of solution	Observed transmission	Calculated transmission	Calculated transmission ratio, 0.99
0.65.....	0.75	39.5	39.7	39.9
0.85.....	.47	56.4	56.5	56.7
1.00.....	.42	60.0	60.2	60.3
1.2.....	.34	66.7	66.5	66.5
1.4.....	.28	71.8	71.6	71.6

Since the original calibration was supposedly based on an average eye corresponding to a characteristic ratio of 1.00, the degree of concordance obtained should be judged, perhaps, by comparison of columns 3 and 5, but as absolute calibrations of the solution columns 3 and 4 should be taken, and within the uncertainty of the present determinations made by the small number of observers mentioned, the equation from which column 4 is calculated ($\log_{10} T = -0.539C^{1.03}$) is correct.

The last four of these solutions were used in measuring the lamps. The 75 per cent concentration is very much bluer than is required by any of the voltages at which these lamps were measured in the comparison between different laboratories, and it was included only for a more complete comparison with the calibration curve and with the previous measurements by a larger number of observers. (See Fig. 5).

The results on the lamps are best shown, perhaps, by comparison with those calculated from the Middlekauff-Skogland equations,¹⁴ which are based on measurements made with the standard Lummer-Brodhun photometer. The following table shows the percentage departure of the values found at the different efficiencies from the calculated values for each lamp. These measurements are equivalent to the direct comparison of lamps at the different efficiencies with a 4 wpc carbon lamp. The differences given are positive when the flicker values are lower.

TABLE 5
Differences Between Standard Lummer-Brodhun and Flicker Photometer Values

Approximate wpc.....	1.4	1.2	1.0	0.85
Lamp 1.....	+0.15	+0.5	+1.45
Lamp 2.....	-.05	+.35	+1.7	+1.35

¹⁴ This Bulletin, 11, p. 483, Sci. Paper No. 238; Trans. Illum. Eng. Soc., 9, p. 734, 1914.

The large departure at 1 wpc is quite inconsistent with all other measurements, and unfortunately the other lamp was not measured at this efficiency. The course of the variation probably should be considered independently of this point. The color match obtainable with this solution is far from being perfect, but the variations arising from this cause are not sufficient to explain the departure of this point from the other measurements. The significance of these measurements can best be seen by referring to the more recent paper by Middlekauff and Skogland, which includes them with results from other laboratories. It may be said that they agree very well with values obtained by Ives using a luminosity scale based on the flicker photometer but fall considerably below all the results obtained with the usual Lummer-Brodhun photometer. The measurements already reported on the 75 per cent concentration of the solution indicate that this difference probably will continue to increase as comparisons are carried to higher efficiencies.

VI. CONCLUSION

1. EFFECT OF INDIVIDUAL CHARACTERISTICS

The preceding results emphasize the fact that for accurate heterochromatic measurements a systematic choice of observers is essential. The system proposed by Ives and Kingsbury appears to be practical and reliable at least for color differences of the type dealt with in these tests. The average eye established is represented by a value of approximately 0.99 for the ratio of the transmission of the yellow test solution to that of the blue solution under the specified conditions. The agreement with the original ratio assigned (1.00) is very good, especially since the latter was largely based on an indirect derivation of the average eye. It is suggested, however, that if the lack of symmetry in the distribution of these observers (as indicated by Fig. 2) is found to persist when larger numbers of observers are included greater reproducibility of the normal ratio might be obtained by choosing as the normal not the average value but the median, which in this case is approximately 0.98.

The differences in observed values arising from individual peculiarities of course increase as the color difference increases. When comparisons are made directly with a 4 wpc carbon lamp, a difference of 1 per cent in the characteristic ratios of observers should result in approximately the following differences in observed

candlepower of a tungsten lamp at the various specific consumptions given:

Watts per candle	Per cent candle-power difference
3.1	0.0
1.4	.1
1.0	.2
.75	.3
.6	.4

When plotted in terms of lumens per watt, these data give nearly a straight line.

Fig. 7 shows the percentage deviation from normal values as a function of the characteristic ratio for the several color differences

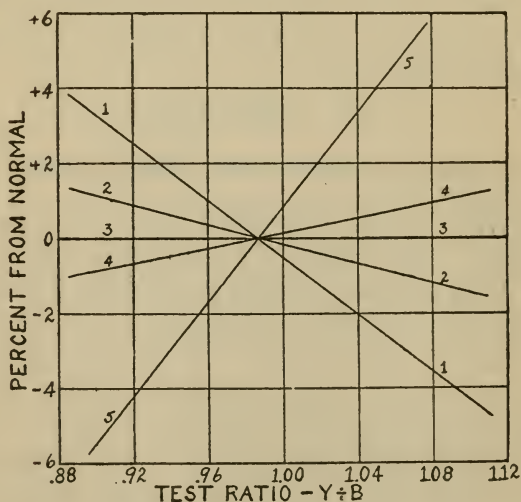


FIG. 7.—Deviation from normal values corresponding to different test ratios when the light indicated is compared with that of a 4 wpc carbon lamp

1. 0.75 concentration blue solution with carbon lamp, equivalent to tungsten lamp at about 0.65 wpc. The curve for the blue test solution practically coincides with this. 2. 1.2 wpc tungsten. 3. Color match, 4 wpc carbon or 3.1 wpc tungsten. 4. Pentane lamp. 5. Yellow test solution

indicated. Thus, an observer whose characteristic ratio $\frac{Y}{B}$ is 0.90, in measuring the candlepower of a 1.2 wpc tungsten lamp against a 4 wpc carbon lamp (3.1 wpc tungsten) would assign to the tungsten lamp a value 1.2 per cent too high. To a 0.65 wpc tungsten lamp he would assign a value 3.5 per cent too high. Of the 114 observers 19 (17 per cent) would have obtained values for a 1.2 wpc tungsten lamp differing 1 per cent or more from the

normal value; 70 (61 per cent) would have departed 1 per cent or more from the normal value in measuring a 0.65 wpc tungsten lamp.

For measurements on color differences of this type it is not necessary to have a group of observers whose average ratio is normal; any observers may be used and their results corrected to the normal by use of curves similar to Fig. 7. These curves can be summarized by the equation

$$I_n = \frac{I}{1 + m(R - R_n)}$$

where I_n is the normal value of a photometric quantity (such as candlepower, illumination, or transmission), R_n is the normal characteristic or "test" ratio, while I and R are the values of these quantities found by a particular observer. m is an empirical constant, which may be either positive or negative and which depends on the color difference involved in the observation. It is the slope of the curves in Fig. 7. There may be slight systematic errors in the corrected values; that is, some observers may be always off the curve in the same direction, and consequently a number of observers should be used for highly accurate results. Not enough repeated measurements are at hand to show with certainty the average magnitude of these systematic errors, but they are certainly small in most cases.

On the average, measurements made with an equality-of-brightness photometer will show practically the same differences due to individual characteristics as those made with the flicker, but the erratic variations are often so great as to overshadow these systematic differences. It has been remarked that some observers develop the ability to make very consistent settings on the equality photometer. In the majority of cases such observers have been found to read close to the value indicated by their test ratio, but this is not always the case.

2. COMPARISON OF FLICKER AND EQUALITY-OF-BRIGHTNESS PHOTOMETERS

With regard to certainty of measurement the flicker photometer shows a decided advantage even with small color differences. With more experienced observers, specially selected, this advantage would probably be materially reduced, but would not be entirely lost, because even when an observer makes consistent settings on the equality photometer the relation of his settings to those of the normal observer is uncertain.

Trained observers are needed with either photometer, but with the flicker any observer of fair ability can make definite sets even with large color differences, whereas on the Lummer-Brodhun photometer it is only the exceptional observer who can do so. Extensive investigations at the English National Physical Laboratory¹⁵ indicate that the final certainty of results is not increased by using the laborious "cascade" or step-by-step method of measurements to avoid sets with large color difference. Little is gained by such a procedure, unless the results of successive steps are agreed upon and made a practically independent standard for future use. This is the tendency of the present practice, and it appears that by this method fairly satisfactorily standards of successively higher and higher temperature may eventually be agreed upon by interlaboratory and international comparisons. No one can say with what degree of accuracy the values of these standards can be reproduced from the fundamental standards a few years hence, and of course, this method applies only to those color differences which can be thus built up step by step with concrete standards to preserve the values at each step.

The flicker photometer, on the other hand, affords a means of relatively precise comparison between lights of all degrees of color difference and makes possible the use of test readings for which average values, which should be highly reproducible, can be established.

In regard to relative results there appears to be no room for doubt that for sources having relatively high intensity at the blue end of the spectrum the values given by the flicker photometer as here used depart appreciably from those obtained with the Lummer-Brodhun as used in common practice, the difference probably being of the order of 3 per cent at the higher efficiencies reached by the present gas-filled lamps. It is, however, hardly proper to assume that the results obtained by either photometer are "right" and anything different is "wrong." The equality-of-brightness method of measurement is undoubtedly more closely related to the way in which the light is used, but it is by no means established that that method correctly indicates the relative usefulness of two kinds of light. It must be recognized that there is no one definite "correct" ratio between the intensities of two lights of different color. The relative candlepowers assigned to a carbon and a tungsten lamp, for example, depend to some extent

¹⁵ Paterson and Dudding, *Proc. Phys. Soc., London*, 27, p. 263, 1915; and *Phil. Mag.* (6), 30, p. 63, 1915.

on the conditions under which the measurements are made. The specification of conditions of measurement must be more or less arbitrary, and the results obtained can not be expected to be an exact indication of the value of different kinds of light under different conditions. Before we shall know much about the relative usefulness of different kinds of light much more experimental work must be done. An important prerequisite for such investigations or any others involving the comparison of the intensity of lights of very different color is a method which will enable different experimenters to make consistent measurements of the quantity which must serve as a basis for the comparison of their results. The usual equality-of-brightness method of comparison certainly does not fulfill this requirement. The flicker photometer at present furnishes the most promising method available.

For the standardizing laboratory, which is expected to reproduce results after the lapse of years when the observers available may be entirely different, the flicker photometer (used under definitely specified conditions), with a systematic method of determining the relation of each observer to the normal, promises to give to heterochromatic photometry a certainty which has appeared quite unattainable with other instruments. Besides giving this probable increased certainty in future reproduction of values, it reduces considerably the labor necessary to attain a given accuracy at the present time. Comparison of actual tests made in the routine work of the laboratory shows that even with relatively small color differences a given accuracy of reproduction of results requires several times as many measurements with the equality-of-brightness or the contrast photometer as with the flicker.

The authors are deeply indebted for the cordial assistance of many associates which has made possible the collection of the data presented. Particular acknowledgement is due to A. H. Taylor and E. M. Baker, who assisted in all of the work.

WASHINGTON, January 15, 1916.



